Claims

[c1]   1. A URL resolution system for resolving Universal Resource Locators (URLs), the URL resolution system comprising:

a website crawler for crawling a website and for locating script code which is used to dynamically create at least one script URL; and

a script URL resolution component for causing examination of the script code located during the crawling to obtain the script URL.

[c2]   2. The URL resolution system as claimed in claim 1 wherein the website includes one or more web pages, and

the website crawler crawls individual web pages associated with websites, and has a crawling controller for controlling the website crawler.

[c3]   3. The URL resolution system as claimed in claim 2 wherein the website crawler has a script code detector for determining if a web page uses script code to dynamically create at least one script URL.

[c4]   4. The URL resolution system as claimed in claim 3 wherein the script code detector has a notification generating function for generating a notification when the script code detector locates a web page that uses script code to dynamically create at least one script URL.

[c5]   5. The URL resolution system as claimed in claim 2 wherein the crawling controller receives results of script code examination from the script URL resolution component, and controls the website crawler based on the examination results.

[c6]   6. The URL resolution system as claimed in claim 5 wherein the examination results include the script URL when the script code examination is successful, and the crawling controller controls the web crawler to crawl a web page identified by the script URL.

[c7]   7. The URL resolution system as claimed in claim 6 wherein the crawling controller controls the website crawler to crawl multiple web pages in parallel.

[c8]   8. The URL resolution system as claimed in claim 6 wherein the crawling

controller controls the website crawler to crawl the web page identified by the script URL immediately.

[c9] 9. The URL resolution system as claimed in claim 6 wherein the crawling controller controls the website crawler to queue the web page identified by the script URL for crawling at a later time.

[c10] 10. The URL resolution system as claimed in claim 1 wherein the script URL resolution component causes execution of the script code to obtain the script URL.

[c11] 11. The URL resolution system as claimed in claim 10 wherein the website includes one or more web pages, the script code has a specific part that is used to create the script URL, and the script URL resolution component comprises: a web page loading controller for instructing a web page examiner to load the web page located by the website crawler; and a script code execution controller for instructing the web page examiner to execute the specific part of the script code used in the loaded web page to obtain the script URL.

[c12] 12. The URL resolution system as claimed in claim 11 wherein the URL resolution system includes the web page examiner.

[c13] 13. The URL resolution system as claimed in claim 11 wherein the script code execution controller uses an execution function of the web page examiner to execute the specific part of the script code.

[c14] 14. The URL resolution system as claimed in claim 11 wherein the website crawler has a script code detector for determining if a web page uses script code to dynamically create at least one script URL, the script code detector generates a notification when it locates a web page that uses script code to dynamically create at least one script URL; and the web page loading controller controls loading of the located web page in response to the notification received from the website crawler.

[c15]
15. The URL resolution system as claimed in claim 11 wherein

the script URL resolution component outputs an execution result including the script URL when the execution of the script code is successful; and

the website crawler performs crawling of a web page identified by the script URL.

[c16]   16. The URL resolution system as claimed in claim 1 wherein
the script URL resolution component outputs an examination result including a failure result when the examination of the script code fails.

[c17]   17. The URL resolution system as claimed in claim 1 wherein the URL resolution system further comprises a presentation unit to present the examination result to a user.

[c18]   18. The URL resolution system as claimed in claim 1 wherein the script URL resolution component is provided as a part of the URL resolution system.

[c19]   19. The URL resolution system as claimed in claim 18 wherein the script URL resolution component is provided as a part of the website crawler.

[c20]   20. The URL resolution system as claimed in claim 13 wherein the website crawler includes the web page examiner.

[c21]   21. The URL resolution system as claimed in claim 1 wherein the website has one or more web pages, and the script URL resolution component is a script URL gatherer for locating each URL contained in any of the web pages of the website and causing examination of a web page identified by each URL to resolve script code contained in the web page to obtain any script URL created by the script code.

[c22]   22. The URL resolution system as claimed in claim 21 further comprising an advanced web page examiner having:
a web page loader for loading a web page identified by a URL received from the script URL gatherer; and
a script code examiner for examining the loaded web page to resolve any script URL that is created by script code in the loaded web page.

[c23]   23. The URL resolution system as claimed in claim 22 wherein the script code

examiner executes script code found in the loaded web page, and returns the execution result to the script URL gatherer.

[c24]   24. The URL resolution system as claimed in claim 22 wherein the advanced web page examiner is provided as a part of the URL resolution system.

[c25]   25. The URL resolution system as claimed in claim 22 wherein
the website crawler further comprises a script code detector for detecting a web page that uses script code to create at least one script URL; and
the script URL gatherer sends to the advanced web page examiner a URL of the web page detected by the script code detector.

[c26]   26. A method for resolving Universal Resource Locators (URLs), the method comprising steps of:
locating script code which creates at least one script URL while crawling a website; and
examining the script code to obtain the script URL from the examination result.

[c27]   27. The method as claimed in claim 26 wherein the locating step locates the script code which creates at least one script URL by examining script code found in the website.

[c28]   28. The method as claimed in claim 26 wherein the examination step obtains the script URL by executing the script code.

[c29]   29. The method as claimed in claim 26 wherein a website has one or more web pages;
the locating step locates a web page that uses script code to dynamically create at least one script URL, the script code having a specific part that is used for the creation of the script URL; and
the examination step comprises steps of:
loading the located web page; and
executing the specific part of the script code in the loaded web page to resolve the script URL.

[c30]
30. The method as claimed in claim 29 wherein the loading step comprises a

step of instructing a web page examiner to load the located web page.

[c31]   31. The method as claimed in claim 29 wherein the execution step comprises a step of instructing a web page examiner to execute the specific part of the script code.

[c32]   32. The method as claimed in claim 29 further comprising a step of continuing crawling of a web page identified by the script URL.

[c33]   33. The method as claimed in claim 32 wherein the continuing step comprises a step of crawling the web page identified by the script URL immediately.

[c34]   34. The method as claimed in claim 32 wherein the continuing step comprises a step of crawling the web page identified by the script URL in parallel to crawling of other web page.

[c35]   35. The method as claimed in claim 32 wherein the continuing step comprises steps of
queuing the web page identified by the script URL; and
crawling the queued web page at a later time.

[c36]   36. The method as claimed in claim 26 wherein a website has one or more web pages and the locating step comprises steps of:
finding a URL in the web pages; and
examining a web page identified by the URL to locate script code in the web page identified by the URL.

[c37]   37. The method as claimed in claim 36 further comprising a step of selecting a web page that contains script code that is used to dynamically create at least one script URL, and wherein the examining step examines the selected web page.

[c38]   38. The method as claimed in claim 26 further comprising steps of:
obtaining examination results including the script URL when the examination step is successful and a failure result when the examination step fails to obtain the script URL; and
presenting to a user the examination result including the script URL and/or the

failure result.

[c39]   39. A computer readable medium storing the instructions and/or statements for use in the execution in a computer of a method for resolving Universal Resource Locators (URLs), the method comprising steps of:

locating script code which creates at least one script URL while crawling a website; and

examining the script code to obtain the script URL from the examination result.

[c40]   40. Electronic signals for use in the execution in a computer of a method for resolving Universal Resource Locators (URLs), the method comprising steps of:

locating script code which creates at least one script URL while crawling a website; and

examining the script code to obtain the script URL from the examination result.

[c41]   41. A computer program product for use in the execution in a computer of a method for resolving Universal Resource Locators (URLs), the computer program product comprising:

a module for locating script code which creates at least one script URL while crawling a website; and

a module for examining the script code to obtain the script URL from the examination result.